BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 7: e112126
doi: 10.3897/biss.7.112126

OPEN ACCESS

Conference Abstract

Can Biodiversity Data Scientists Document Volunteer and Professional Collaborations and Contributions in the Biodiversity Data Enterprise?

Robert D. Stevenson[‡], Elizabeth R. Ellwood[§], Peter Brenton[|], Paul Kenneth John Flemons[¶], Jeff Gerbracht[#], Wesley M. Hochachka[#], Scott Loarie[¤], Carrie Seltzer[¤]

‡ University of Massachusetts Boston, Boston, MA, United States of America
§ Natural History Museum of Los Angeles County, Los Angeles, United States of America
| Atlas of Living Australia, Canberra, Australia
¶ Australian Museum, Sydney, Australia
# Cornell Lab of Ornithology, Ithaca, NY, United States of America
¤ iNaturalist, San Rafael, CA, United States of America

Corresponding author: Robert D. Stevenson (robert.stevenson@umb.edu)

Received: 04 Sep 2023 | Published: 06 Sep 2023

Citation: Stevenson RD, Ellwood ER, Brenton P, Flemons PKJ, Gerbracht J, Hochachka WM, Loarie S, Seltzer C (2023) Can Biodiversity Data Scientists Document Volunteer and Professional Collaborations and Contributions in the Biodiversity Data Enterprise? Biodiversity Information Science and Standards 7: e112126. https://doi.org/10.3897/biss.7.112126

Abstract

The collection, archiving and use of biodiversity data depend on a network of pipelines herein called the Biodiversity Data Enterprise (BDE) and best understood globally through the work of the Global Biodiversity Information Facility (GBIF). Efforts to sustain and grow the BDE require information about the data pipeline and the infrastructure that supports it. A host of metrics from GBIF, including institutional participation (member countries, institutional contributors, data publishers), biodiversity coverage (occurrence records, species, geographic extent, data sets) and data usage (records downloaded, published papers using the data) (Miller 2021), document the rapid growth and successes of the BDE (GBIF Secretariat 2022). Heberling et al. (2021) make a convincing case that the data integration process is working.

The Biodiversity Information Standards' (TDWG) Basis of Record term provides information about the underlying infrastructure. It categorizes the kinds of processes*[1] that

This is an open access article distributed under the terms of the CC0 Public Domain Dedication.

teams undertake to capture biodiversity information and GBIF quantifies their contributions*[2] (Table 1). Currently 83.4% of observations come from human observations, of which 63% are of birds. Museum preserved specimens account for 9.5% of records. In both cases, a combination of volunteers (who make observations, collect specimens, digitize specimens, transcribe specimen labels) and professionals work together to make records available.

Table 1.

Data Categories in GBIF as of June 30, 2023.

Category	Number of Contributions	Fraction of Contributions
Observation	23,399,199	0.010
Machine observation	16,717,275	0.007
Human observation	1,971,657,293	0.834
Material sample	54,790,163	0.023
Material citation	3,180,597	0.001
Preserved specimen	224,583,775	0.095
Fossil specimen	10,148,965	0.004
Living specimen	1,997,262	0.001
Occurrence	56,298,292	0.024
Total	2,362,772,821	1

To better understand how the BDE is working, we suggest that it would be of value to know the number of contributions and contributors and their hours of engagement for each data set. This can help the community address questions such as, "How many volunteers do we need to document birds in a given area?" or "How much professional support is required to run a camera trap network?" For example, millions of observations were made by tens of thousands of observers in two recent BioBlitz events, one called Big Day, focusing on birds, sponsored by the Cornell Laboratory of Ornithology and the other called the City Nature Challenge, addressing all taxa, sponsored jointly by the California Academy of Sciences and the Natural History Musuems of Los Angeles County (Table 2). In our presentation we will suggest approaches to deriving metrics that could be used to document the collaborations and contribution of volunteers and staff using examples from both Human Observation (eBird, iNaturalist) and Preserved Specimen (DigiVol, Notes from Nature) record types. The goal of the exercise is to start a conversation about how such metrics can further the development of the BDE.

Table 2.

Examples of the outcomes, numbers of permanent staff and participants collaborating (with biodiversity outcomes) on two citizen science bioblitzes in 2023: Most of the outcome data are from the two links eBird's Big Day and iNaturalist's City Nature Challenge (as of August 14, 2023 for iNaturalist). Other data sources are in endnotes.

Event Characteristics	**Big Day**	**City Nature Challenge**
Sponsoring organizations	Cornell Laboratory of Ornithology	Natural History Museums of Los Angeles County & California Academy of Sciences
Collection platform	eBird	iNaturalist
Collection time frame	13-May-23	28 April -1 May 1, 2023
Staff involved	~30*[3]	~20*[4]
Local organizers	>150*[5]	>800*[6]
Expert reviewers	~2,222*[7]	-
ID contributors	-	19,408
Participants	58,756	68,855
Taxonomic scope	Birds	All taxa
Biodiversity observations (millions)	3.2	1.87
Species obsserved	7,636	58,088
Countries involved	199	46

Keywords

data science, community, TDWG, GBIF, iNaturalist, eBird

Presenting author

Paul Kenneth John Flemons

Presented at

TDWG 2023

Conflicts of interest

The authors have declared that no competing interests exist.

References

- GBIF Secretariat (2022) GBIF Science Review 2021. URL: https://www.gbif.org/document/57InivSC2VzMFWbTJbqOsD/gbif-science-review-2021
- Heberling JM, Miller J, Noesgaard D, Weingart S, Schigel D (2021) Data integration enables global biodiversity synthesis. Proceedings of the National Academy of Sciences 118 (6). https://doi.org/10.1073/pnas.2018093118
- Miller JM (2021) GBIF relevance for science and policy. In: GBIF (Ed.) Impact and Action: A Virtual Science-Policy Symposium. 09/12/2021. GBIF, Copenhagen [In English]. URL: https://www.gbif.org/event/5F0JBfJgGORf3pzyusksKa/impact-and-action-a-virtual-science-policy-symposium

Endnotes

*1 GBIF Secretariat (2022) Biodiversity Data Use. Version B34f741, 2022-03-30 08:40:29 UTC

*2 GBIF (2023). GBIF occurrence tab. https://www.gbif.org/occurrence/search. Accessed on: 2023-7-01.

*3 See https://www.gbif.org/dataset/4fa7b334-ce0d-4e88-aaae-2e0c138d049e

*4 The Californai Academy of Sciences and the Natural History Museums of Los Angles County Museum staff team up with the iNaturalist organization to run the event

*5 Wesley Hochachka's estimation

*6 There were 452 cities involved (see https://www.inaturalist.org/posts/79679-results-of-the-2023-city-nature-challenge). The number of organizers varied from city to city from a large team of 10 or more to a just one or two people. We think 800 is a conservative number.

*7 https://ebird.org/news/2022-year-in-review